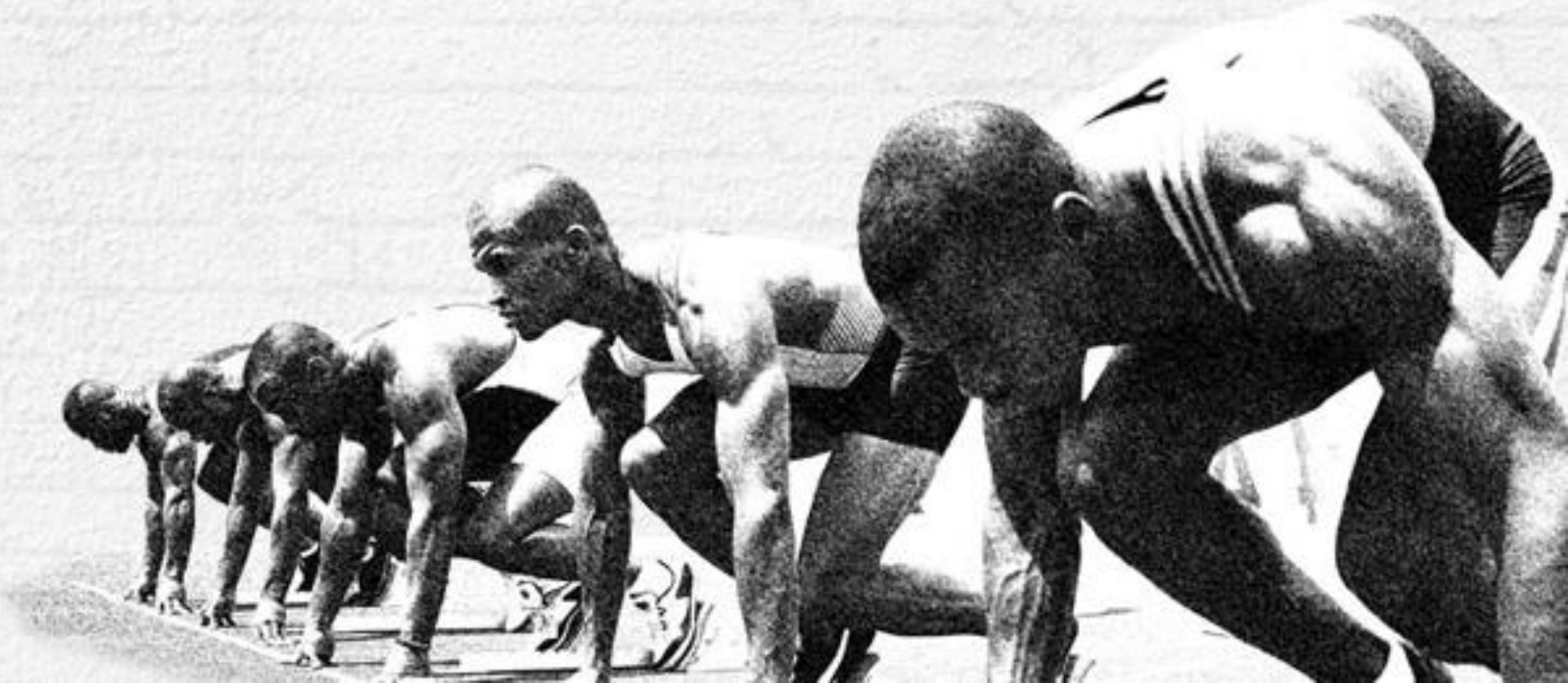
Using Apache Kafka to replicate data from Postgres to SQL Server

Marcelo Adade

Principal Database Consultant

Pythian



DATA PLATFORM VIRTUAL SUMMIT 2022

Diamond Knowledge Partner Microsoft

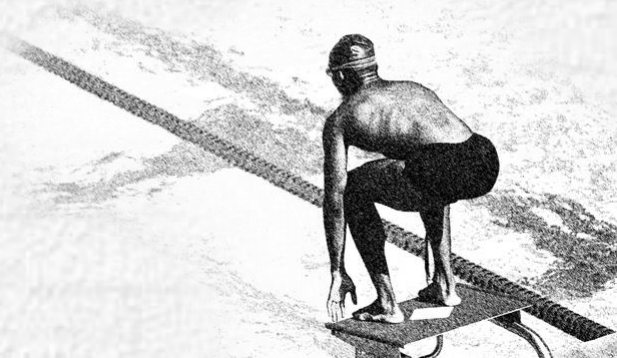
Marcelo Adade

- **Microsoft Data Platform Internal Principal Consultant at Pythian (Ottawa)**
- **Database Consultant and Instructor at DBBITS (Sao Paulo)**
- **Microsoft Data Platform MVP**
- https://linktr.ee/marceloadade
- https://pythian.com/careers/

Pythian
love your data®

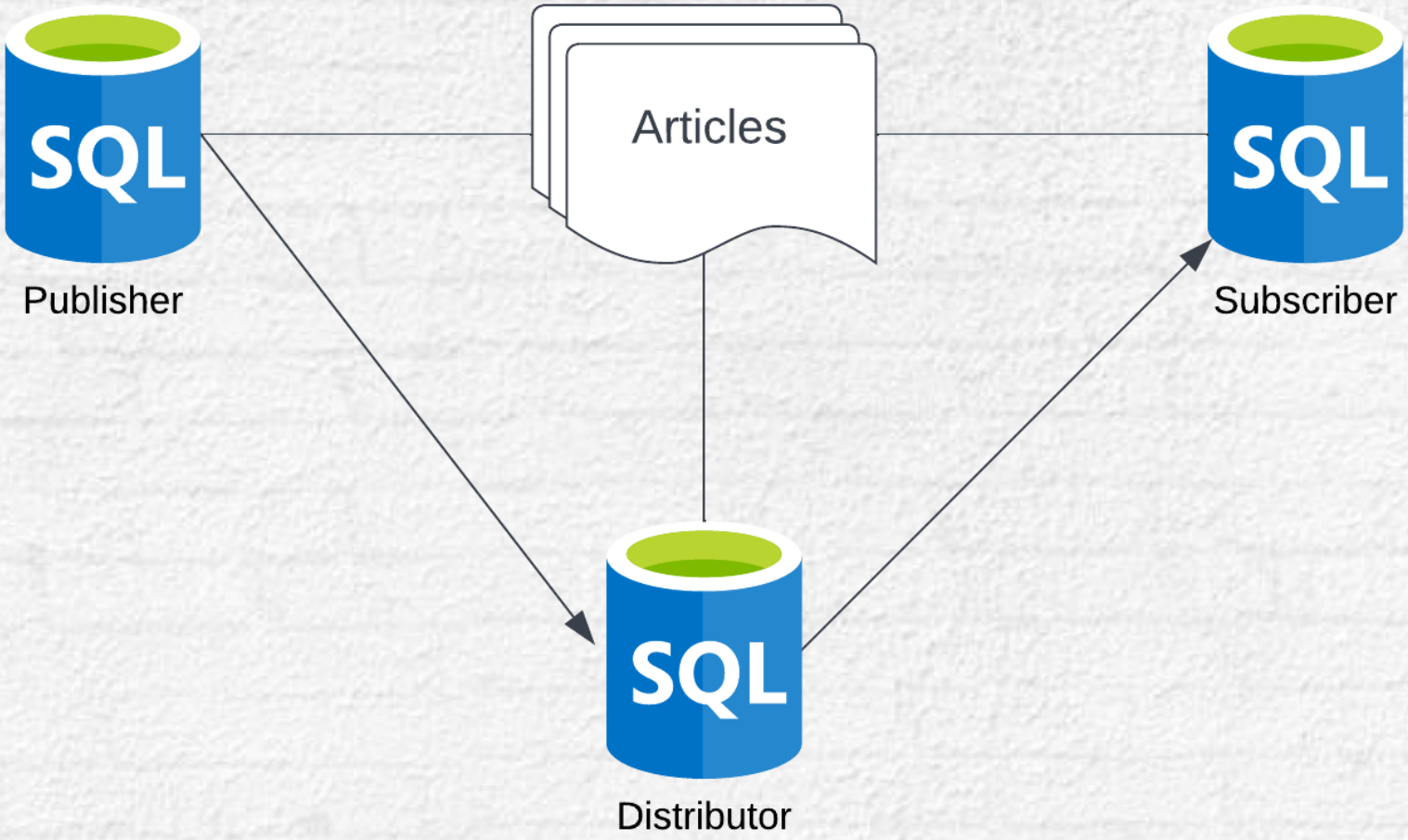
PPT, demos and scripts?

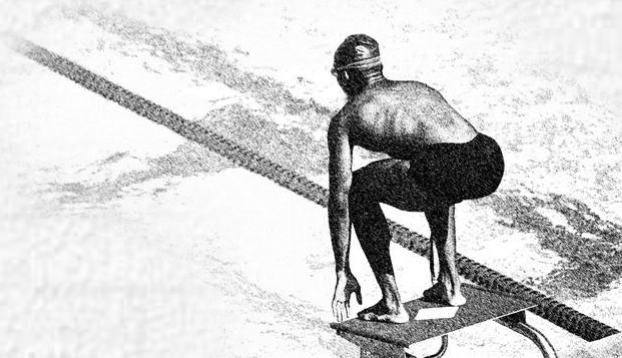
- https://bit.ly/dps2022-kafkapsql



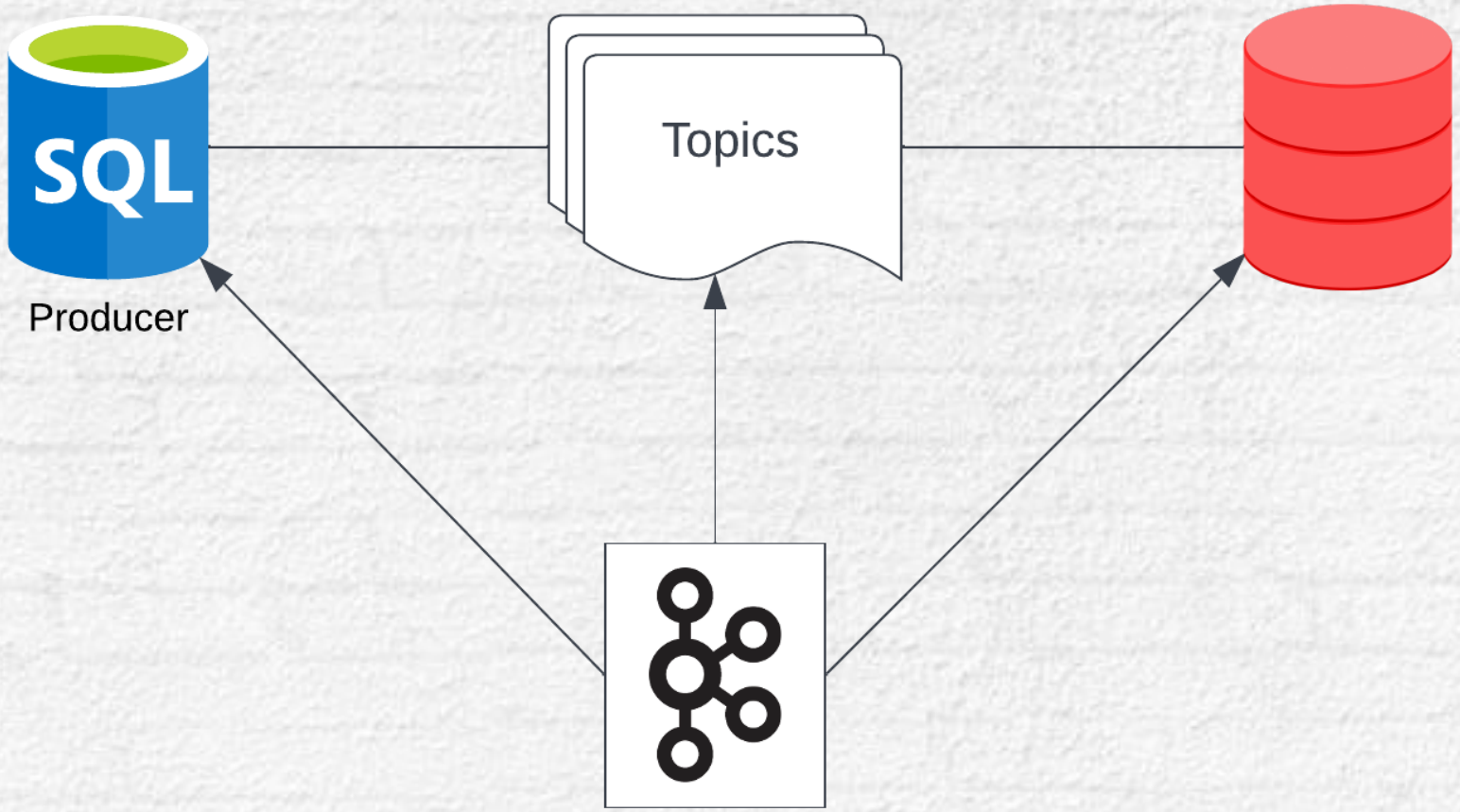
THE SCENARIO

Native Replication



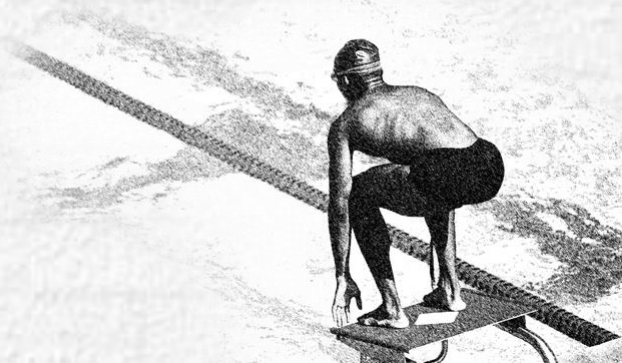


Heterogeneous replication



The problem

- No official support
- Only Oracle was available
- No heterogeneous replication from the Pgsql side
- A third-party tool would be somewhat expensive



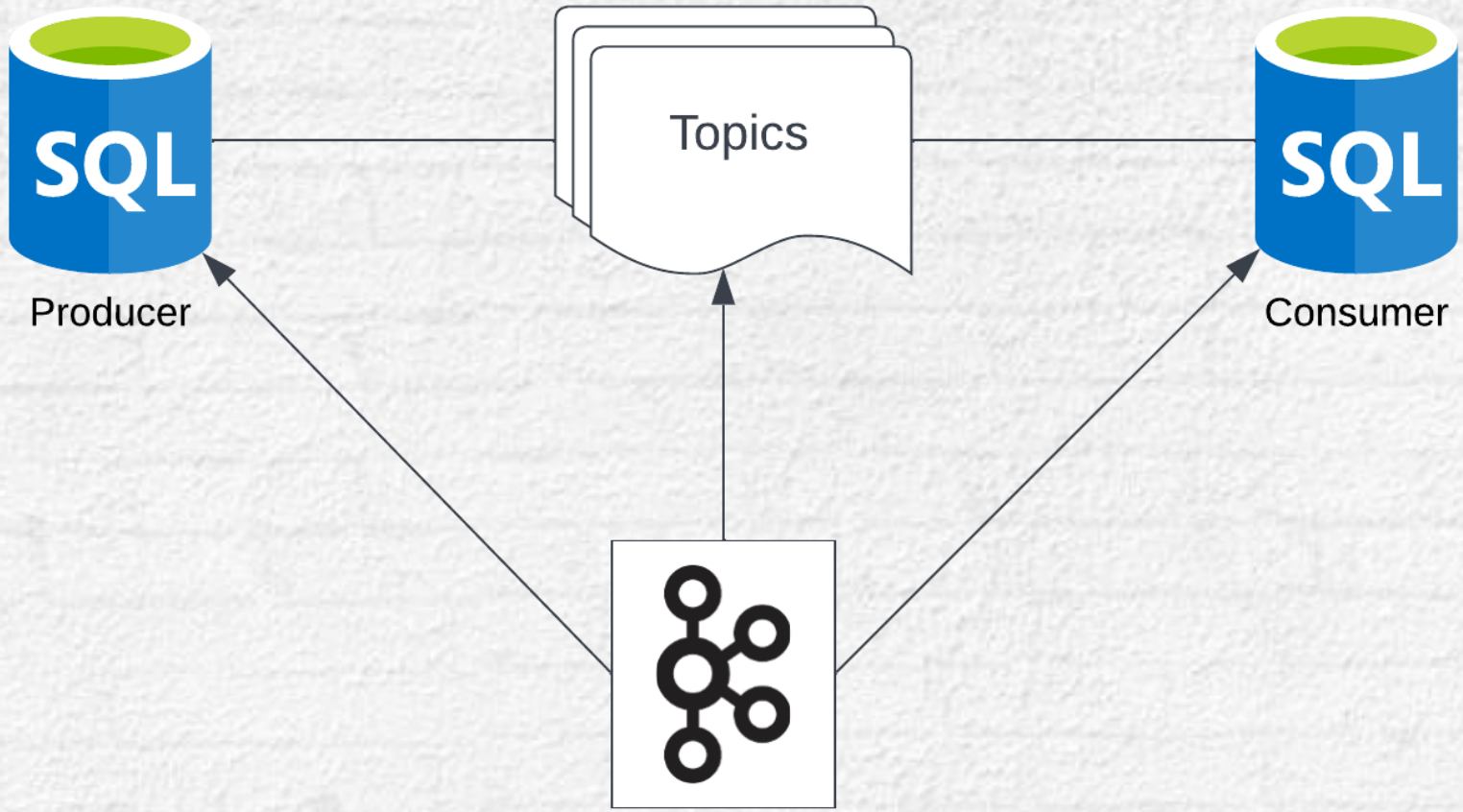
Possible solutions

Qlik

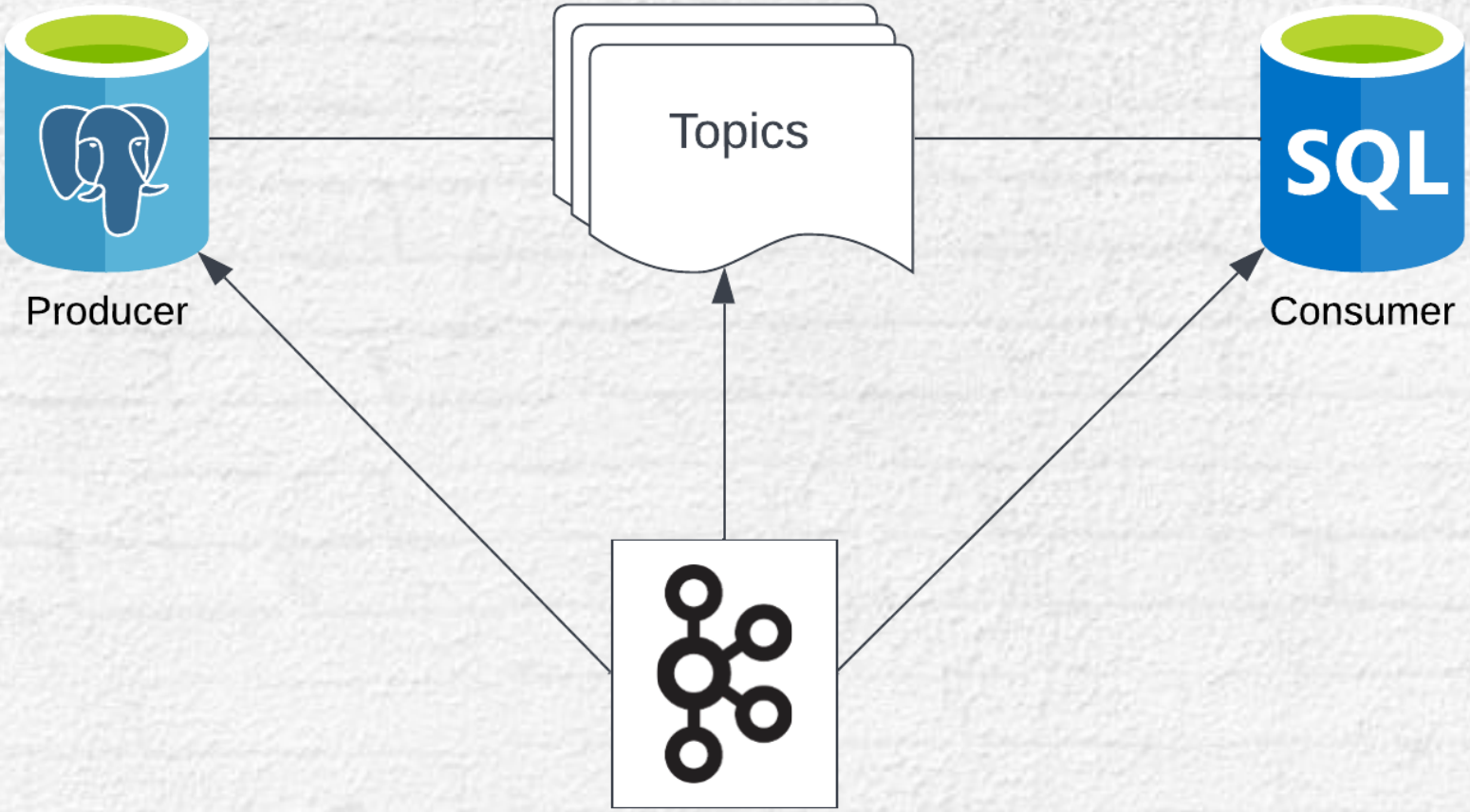
CLOUDBASIX

kafka

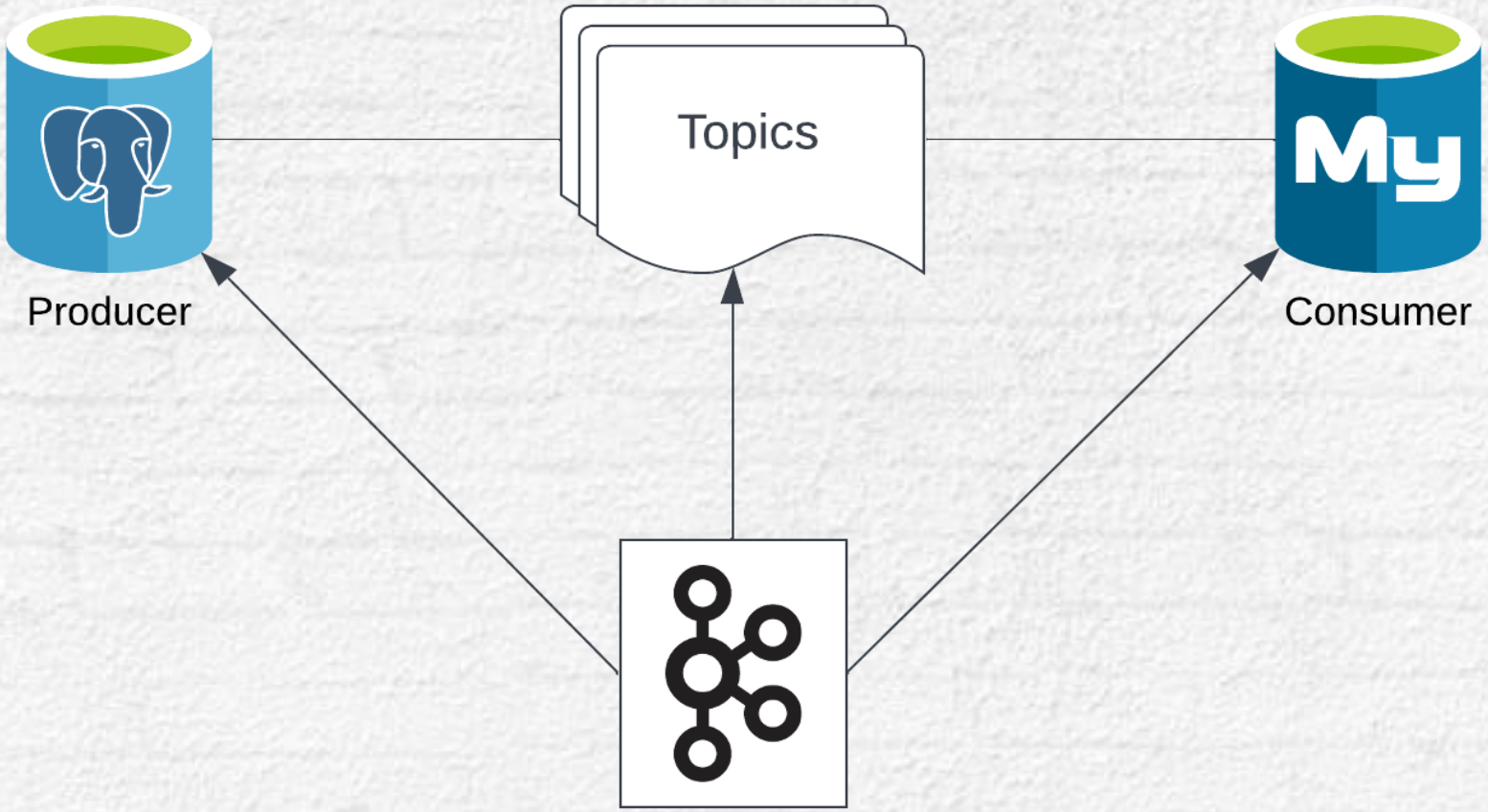
Enter Apache Kafka

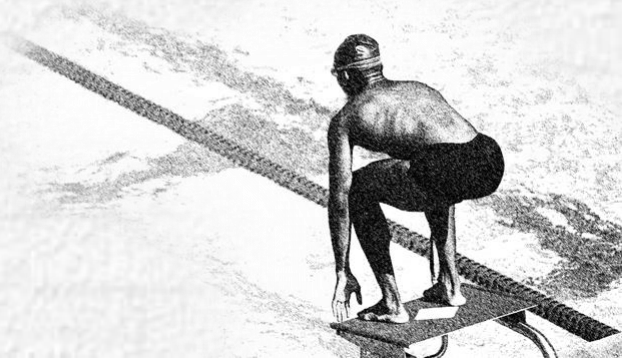


Enter Apache Kafka



Enter Apache Kafka



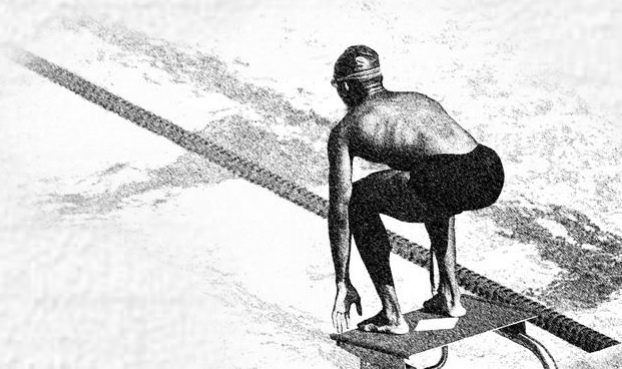


SOLUTION 1

"Manual" Kafka

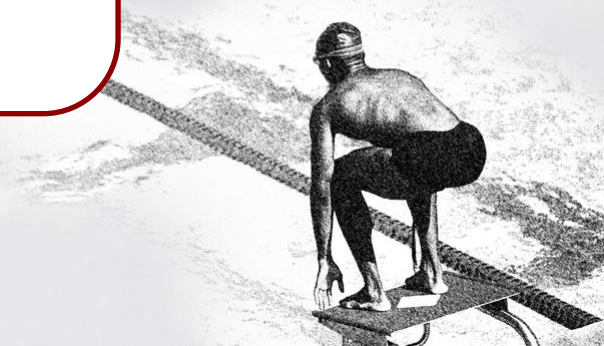
Manual Step-by-step

1. Define a data source or Producer
2. Define a data destination or Consumer
3. Install the Confluent Kafka platform
4. Install the required drivers
5. Create the connectors using the drivers
6. Monitor the Data flow
7. Sit back and relax!



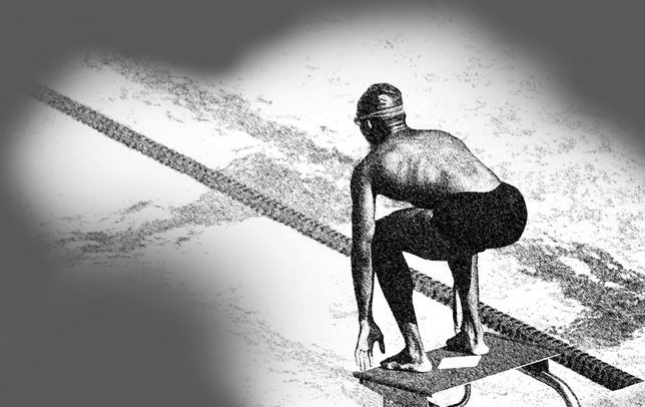
Demo

Setting up the Confluent Platform



SOLUTION 2

Cloud Kafka



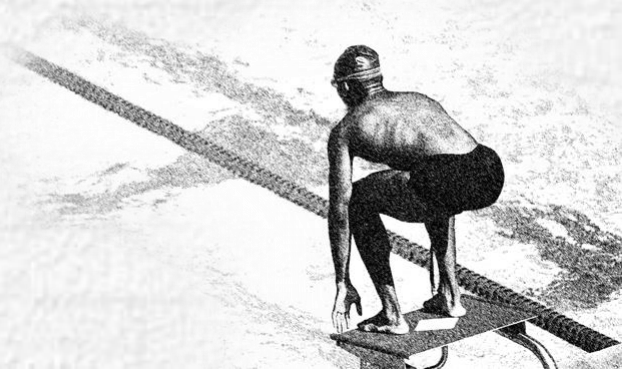
Confluent Cloud

CONFLUENT

Google Cloud

aws

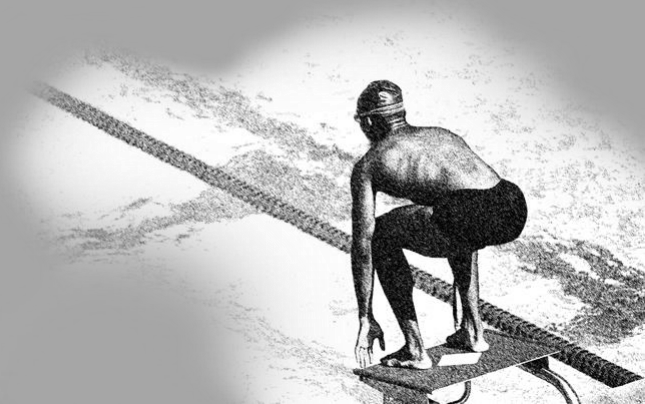
https://www.confluent.io/confluent-cloud/



RECAP

Two solutions for replicating Data thru different Database technologies:

- Using Apache Kafka and setting it up manually
- Using Confluent Cloud within your provider of choice



Special Thanks To

Microsoft

for supporting

DataPlatformGeeks & SQLServerGeeks

Community Initiatives

SQLServerGeeks

Data Platform Summit

DataPlatformGeeks

Thank You

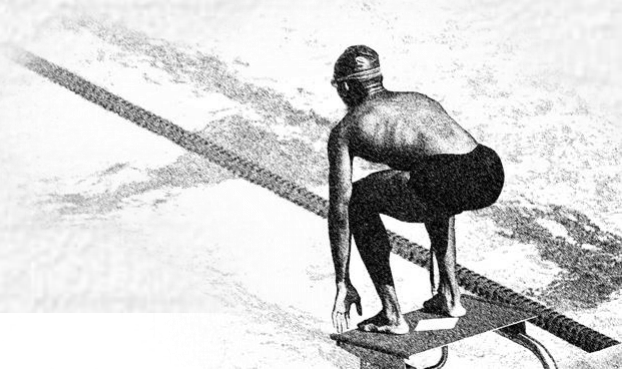
Three Ways to Win Prizes

Post your selfie with hash tag **#DPS2022**

Give Session & Conference Feedback

Visit our Sponsors & Exhibitors

Follow us on Twitter **@TheDataGeeks @DataAISummit**

DATA PLATFORM
VIRTUAL SUMMIT 2022

Diamond Knowledge Partner Microsoft